5. The compactness theorem

Here we prove the *compactness theorem*: If a set of sentences is such that every finite subset of it has a model, then the whole set has a model. The theorem will be an easy consequence of Łoś's theorem on ultraproducts.

Theorem 5.1. (Łoś) *Suppose that $\mathscr{L}$ is a first-order language, $\overline{A} = \langle \overline{A_i} : i \in I\rangle$ is a system of $\mathscr{L}$-structures, F is an ultrafilter on I, and $a \in {}^{\omega}\prod_{i\in I} a_i$. The values of a will be denoted by $a^0, a^1, \ldots$. Let $\pi : \prod_{i\in I} A_i \to \prod_{i\in I} A_i/F$ be the natural mapping, taking each element of $\prod_{i\in I} A_i$ to its equivalence class under $\equiv^A_F$. For each $i \in I$ let $pr_i : \prod_{j\in I} A_j \to A_i$ be defined by setting $pr_i(x) = x_i$ for all $x \in \prod_{i\in I} A_i$. Suppose that φ is any formula of $\mathscr{L}$. Then*

$$\prod_{i\in I} \overline{A_i}/F \models \varphi[\pi\circ a] \text{ iff } \{i \in I : \overline{A_i} \models \varphi[\mathrm{pr}_i \circ a]\} \in F.$$

Proof. Because the situation and the notation are complicated, we are going to give the proof in full. For brevity let $\overline{B} = \prod_{i\in I} \overline{A_i}/F$. First we show

(1) For any term τ, $\tau^{\overline{B}}(\pi\circ a) = [\langle \tau^{\overline{A_i}}(\mathrm{pr}_i\circ a) : i \in I\rangle]_F$.

We prove (1) by induction on τ. For τ a variable v_k,

$$\tau^{\overline{B}}(\pi\circ a) = (\pi\circ a)(k) = [a^k]_F = [\langle a^k_i : i\in I\rangle]_F = [\langle v_k^{\overline{A_i}}(\mathrm{pr}_i\circ a) : i\in I\rangle]_F,$$

as desired. For τ an individual constant $\mathbf{k}$,

$$\mathbf{k}^{\overline{B}}(\pi\circ a) = \mathbf{k}^{\overline{B}} = [\langle \mathbf{k}^{\overline{A_i}} : i\in I\rangle]_F = [\langle \mathbf{k}^{\overline{A_i}}(\mathrm{pr}_i\circ a) : i\in I\rangle]_F.$$

The inductive step:

$$\begin{aligned}
(\mathbf{F}\sigma_0\ldots\sigma_{m-1})^{\overline{B}}(\pi\circ a) &= \mathbf{F}^{\overline{B}}(\sigma_0^{\overline{B}}(\pi\circ a),\ldots,\sigma_{m-1}^{\overline{B}}(\pi\circ a))\\
&= \mathbf{F}^{\overline{B}}([\langle\sigma_0^{\overline{A_i}}(\mathrm{pr}_i\circ a) : i\in I\rangle]_F,\ldots,[\langle\sigma_{m-1}^{\overline{A_i}}(\mathrm{pr}_i\circ a) : i\in I\rangle]_F)\\
&= [\langle \mathbf{F}^{\overline{A_i}}(\sigma_0^{\overline{A_i}}(\mathrm{pr}_i\circ a),\ldots,\sigma_{m-1}^{\overline{A_i}}(\mathrm{pr}_i\circ a)) : i\in I\rangle]_F\\
&= [\langle \tau^{\overline{A_i}}(\mathrm{pr}_i\circ a) : i\in I\rangle]_F,
\end{aligned}$$

as desired.

Now we begin the real proof of the theorem, proceding, of course, by induction on φ. Suppose that φ is $\sigma = \tau$. Then

$$\begin{aligned}
\overline{B} \models (\sigma=\tau)[\pi\circ a] &\text{ iff } \sigma^{\overline{B}}(\pi\circ a) = \tau^{\overline{B}}(\pi\circ a)\\
&\text{ iff } [\langle\sigma^{\overline{A_i}}(\mathrm{pr}_i\circ a) : i\in I\rangle]_F = [\langle\tau^{\overline{A_i}}(\mathrm{pr}_i\circ a) : i\in I\rangle]_F\\
&\text{ iff } \{i\in I : \sigma^{\overline{A_i}}(\mathrm{pr}_i\circ a) = \tau^{\overline{A_i}}(\mathrm{pr}_i\circ a)\} \in F\\
&\text{ iff } \{i\in I : \overline{A_i} \models (\sigma=\tau)[\mathrm{pr}_i\circ a]\} \in F,
\end{aligned}$$

as desired.

Now suppose that φ is $\mathbf{R}\sigma_0 \dots \sigma_{m-1}$. Then

$$\begin{aligned}
\overline{B} \models \varphi[\pi \circ a] &\text{ iff } (\sigma_0^{\overline{B}}(\pi \circ a), \dots, \sigma_{m-1}^{\overline{B}}(\pi \circ a)) \in \mathbf{R}^{\overline{B}} \\
&\text{ iff } ([\langle \sigma_0^{\overline{A_i}}(\mathrm{pr}_i \circ a) : i \in I\rangle]_F, \dots, [\langle \sigma_{m-1}^{\overline{A_i}}(\mathrm{pr}_i \circ a) : i \in I\rangle]_F) \in \mathbf{R}^{\overline{B}} \\
&\text{ iff } \{i \in I : (\sigma_0^{\overline{A_i}}(\mathrm{pr}_i \circ a), \dots, \sigma_{m-1}^{\overline{A_i}}(\mathrm{pr}_i \circ a)) \in \mathbf{R}^{\overline{A_i}}\} \in F \\
&\text{ iff } \{i \in I : \overline{A_i} \models \varphi[\mathrm{pr}_i \circ a]\} \in F,
\end{aligned}$$

as desired.

The inductive step when φ is $\neg\psi$:

$$\begin{aligned}
\overline{B} \models \varphi[\pi \circ a] &\text{ iff not}(\overline{B} \models \varphi[\pi \circ a]) \\
&\text{ iff not}(\{i \in I : \overline{A_i} \models \psi[\mathrm{pr}_i \circ a]\} \in F) \\
&\text{ iff } I \backslash \{i \in I : \overline{A_i} \models \psi[\mathrm{pr}_i \circ a]\} \in F \\
&\text{ iff } \{i \in I : \text{not}(\overline{A_i} \models \psi[\mathrm{pr}_i \circ a])\} \in F \\
&\text{ iff } \{i \in I : \overline{A_i} \models \varphi[\mathrm{pr}_i \circ a]\} \in F,
\end{aligned}$$

as desired.

The induction step for $\wedge$:

$$\begin{aligned}
\overline{B} \models (\psi \wedge \chi)[\pi \circ a] & \overline{B} \models \psi[\pi \circ a]) \text{ and } \overline{B} \models \chi[\pi \circ a] \\
& \{i \in I : \overline{A_i} \models \psi[\mathrm{pr}_i \circ a]\} \in F) \text{ and } \{i \in I : \overline{A_i} \models \chi[\mathrm{pr}_i \circ a]\} \in F \\
& \text{ iff } \{i \in I : \overline{A_i} \models \psi[\mathrm{pr} \circ a] \text{ and } \overline{A_i} \models \chi[\mathrm{pr} \circ a]\} \in F \\
& \text{ iff } \{i \in I : \overline{A_i} \models (\psi \wedge \chi)[\mathrm{pr}_i \circ a]\} \in F,
\end{aligned}$$

as desired.

It remains only to consider φ of the form $\exists v_k \psi$, which is the main case. We do each direction in the desired equivalence separately. First suppose that $\overline{B} \models \varphi[\pi \circ a]$. Choose $u \in \prod_{i \in I} A_i$ such that $\overline{B} \models \psi[(\pi \circ a)^k_{[u]_F}]$. Now $(\pi \circ a)^k_{[u]_F} = \pi \circ a^k_u$, so we can apply the induction hypothesis and get $\{i \in I : \overline{A_i} \models \psi[\mathrm{pr}_i \circ a^k_u]\} \in F$. But for each $i \in I$ we have $(\mathrm{pr}_i \circ a^k_u) = (\mathrm{pr}_i \circ a)^k_{u(i)}$, so

$$\{i \in I : \overline{A_i} \models \psi[\mathrm{pr}_i \circ a^k_u]\} \subseteq \{i \in I : \overline{A_i} \models \varphi[\mathrm{pr}_i \circ a]\},$$

and hence $\{i \in I : \overline{A_i} \models \varphi[\mathrm{pr}_i \circ a]\} \in F$. This finishes half of what we want.

Conversely, suppose that $\{i \in I : \overline{A_i} \models \varphi[\mathrm{pr}_i \circ a]\} \in F$. For each i in this set pick $u(i) \in A_i$ such that $\overline{A_i} \models \psi[(\mathrm{pr}_i \circ a)^k_{u(i)}]$ (using the axiom of choice). For other i's in I let $u(i)$ be any old element of A_i, just to fill out u to make it a member of $\prod_{i \in I} A_i$. Now $(\mathrm{pr}_i \circ a)^k_{u(i)} = \mathrm{pr}_i \circ a^k_u$. Thus $\{i \in I : \overline{A_1} \models \psi[\mathrm{pr}_i \circ a^k_u]\} \in F$. By the inductive hypothesis it follows that $\overline{B} \models \psi[\pi \circ a^k_u]$. But $\pi \circ a^k_u = (\pi \circ a)^k_{[u]_F}$, so we finally get $\overline{B} \models \varphi[\pi \circ a]$. $\square$

Corollary 5.2. *Suppose that $\mathscr{L}$ is a first-order language, $\overline{A} = \langle \overline{A_i} : i \in I\rangle$ is a system of $\mathscr{L}$-structures, and F is an ultrafilter on I. Suppose that φ is any sentence of $\mathscr{L}$. Then*

$$\prod_{i\in I} \overline{A_i}/F \models \varphi \text{ iff } \{i \in I : \overline{A_i} \models \varphi\} \in F. \qquad \square$$

The compactness theorem is now easy to prove:

Theorem 5.3. (The Compactness Theorem) *Suppose that Γ is a set of sentences in a first-order language $\mathscr{L}$, and every finite subset of Γ has a model. Then Γ has a model.*

Proof. For each finite subset Δ of Γ, let $\overline{A}_\Delta$ be a model of Δ. Let $I = \{\Delta \subseteq \Gamma : \Delta$ is finite$\}$. For each $\Delta \in I$, let

$$x_\Delta = \{\Theta \in I : \Delta \subseteq \Theta\}.$$

Then the family $\{x_\Delta : \Delta \in I\}$ of subsets of I has fip. In fact, let $\Delta_0, \ldots, \Delta_{m-1} \in I$. Then clearly

$$\Delta_0 \cup \ldots \cup \Delta_{m-1} \in x_{\Delta_0} \cap \ldots \cap x_{\Delta_{m-1}},$$

as desired. Hence by Theorem 1.14, let F be an ultrafilter on I such that $x_\Delta \in F$ for all $\Delta \in I$. We claim now that $\prod_{\Delta\in I} \overline{A}_\Delta/F$ is a model of Γ. To see this, take any $\varphi \in \Gamma$. Then for any $\Delta \in x_{\{\varphi\}}$ we have $\varphi \in \Delta$, and $\overline{A}_\Delta$ is a model of Δ, hence of φ. Thus $\{\Delta \in I : \overline{A}_\Delta \models \varphi\} \supseteq x_\Delta$, so $\{\Delta \in I : \overline{A}_\Delta \models \varphi\} \in F$. By Corollary 5.2 it follows that $\prod_{\Delta\in I} \overline{A}_\Delta/F \models \varphi$, as desired. $\square$

We now indicate a few applications of the compactness theorem. There are more in the exercises.

Corollary 5.4. *Let $\mathscr{L}$ be a first order language. Then there does not exist a set Γ of sentences of $\mathscr{L}$ such that the models of Γ are exactly the finite $\mathscr{L}$-structures.*

Proof. Suppose that such a set Γ exists. Then let

$$\Theta = \Gamma \cup \left\{ \exists v_0 \ldots \exists v_{m-1} \left(\bigwedge_{i<j<m} \neg(v_i = v_j) \right) : m \in \omega \right\}.$$

Note that for a given m, the sentence $\exists v_0 \ldots \exists v_{m-1}(\bigwedge_{i<j<m} \neg(v_i = v_j)$ holds in a structure $\overline{A}$ iff A has at least m elements. Since there are $\mathscr{L}$-structures of any finite size, it follows that every finite subset of Θ has a model. So Θ itself has a model. But such a model must satisfy the indicated sentences for all $m < \omega$, and so must be infinite, contradiction. $\square$

The argument just given really shows the following stronger result:

Corollary 5.5. *Let $\mathscr{L}$ be a first order language and Γ a set of sentences of $\mathscr{L}$. If Γ has models of arbitrarily large finite size (i.e., if for every $m \in \omega$ Γ has a finite model with at least m elements), then Γ has an infinite model.* $\square$

Corollary 5.6. *If a sentence φ holds in every infinite model of a set Γ of sentences, then there is an $m \in \omega$ such that it holds in every model of Γ with at least m elements.*

Proof. Suppose that the conclusion fails. Note that the hypothesis implies that every model of $\Gamma \cup \{\neg\varphi\}$ is finite. The conclusion failing means that for every $m \in \omega$ there is a model of Γ with at least m elements in which φ fails. So $\Gamma \cup \{\neg\varphi\}$ has models of arbitrarily large finite size, hence by 5.5 it has an infinite model, contradiction. $\square$

Note that Corollary 5.6 can be applied in the theory of fields: let Γ be a collection of sentences saying that the $\mathscr{L}$-structure is a field of characteristic p (p fixed). Thus a sentence holds in every infinite field of characteristic p (p a prime) iff there is an m such that it holds in every field of characteristic p with at least m elements.

Corollary 5.7. *Let $\mathscr{L}$ be the language of ordering. Then there is no set Γ of sentences whose models are exactly the well-ordering structures.*

Proof. Suppose there is such a set. Let us expand the language $\mathscr{L}$ to a new one $\mathscr{L}'$ by adding an infinite sequence $\mathbf{c}_m$, $m \in \omega$, of individual constants. Then consider the following set Θ of sentences: all members of Γ, plus all sentences $\mathbf{c}_{m+1} < \mathbf{c}_m$ for $m \in \omega$. Clearly every finite subset of Θ has a model, so let $\overline{A} = (A, <, a_i)_{i<\omega}$ be a model of Θ itself. (Here a_i is the 0-ary function, i.e., element of A, corresponding to $\mathbf{c}_i$.) Then $a_0 > a_1 > \cdots$; so $\{a_i : i \in \omega\}$ is a nonempty subset of A with no least element, contradiction. $\square$

The applications of the compactness theorem so far are negative: certain things cannot be done in first-order logic. Positive applications are more interesting, but generally more lengthy. We give one example.

Proposition 5.8. *If $(A, <)$ is a partial ordering structure, then there is a relation $\prec$ such that $(A, \prec)$ is a simple ordering structure and $<$ is a subset of $\prec$.*

Proof. First we prove this for A finite. This goes by induction on $|A|$. It is obvious if $|A| = 1$. If we have proved it for all partial ordering structures $(A, <)$ with $|A| = m$, suppose that $(A, <)$ is a partial ordering structure with $|A| = m + 1$. Fix $a_0 \in A$, and let $A' = A\backslash\{a_0\}$ and $<' = < \cap (A' \times A')$. Let $\prec'$ be a simple ordering on A' such that $<'$ is a subset of $\prec'$. Let

$$X = \{u \in A' : \exists v(v < a_0 \text{ and } u \preceq' v)\},$$

and then we define a relation $\prec$ on A by setting, for any $x, y \in A$,

$$x \prec y \text{ iff } \begin{cases} x, y \in A' \text{ and } x \prec' y \\ \text{or } x \in X \text{ and } y = a_0 \\ \text{or } x = a_0 \text{ and } y \in A\backslash(X \cup \{a_0\}). \end{cases}$$

It is straightforward but tedious to check that this gives a simple order extending $<$; we go through the details. Clearly $\prec$ is irreflexive. Suppose now that $x \prec y \prec z$; we want to show that $x \prec z$. If $x, y, z \in A'$ this is obvious. So three cases remain. *Case 1.* $x = a_0$. Then $y \in A\backslash(X \cup \{a_0\})$. Then $y \notin X$, and this implies that $z \neq a_0$. Suppose that $z \in X$. Choose accordingly v such that $v < a_0$ and $z \preceq' v$. Now also $y \prec' z$, so $y \prec' v$, and hence $y \in X$, contradiction. Thus $z \notin X$, so $x \prec z$. *Case 2.* $y = a_0$. Thus $x \in X$ and $z \in A\backslash(X \cup \{a_0\})$.

Since $x \in X$, choose v so that $v < a_0$ and $x \preceq' v$. Obviously $z \neq x$. Suppose that $z \prec' x$. Then $z \prec' v$ and so $z \in X$, contradiction. The only remaining possibility is that $x \prec' z$, hence $x \prec z$. *Case 3.* $z = a_0$. Thus $y \in X$ and $x \prec' y$. It is clear then from the definition of X that also $x \in X$, and so $x \prec' z$.

Next we show that $(A, \prec)$ is a simple ordering structure. Suppose that x and y are distinct elements of A. If both of them are in A', then $x \prec' y$ or $y \prec' x$, hence $x \prec y$ or $y \prec x$. Suppose that one of them, say x, is a_0. Then either $y \in X$, in which case $y \prec x$, or $y \notin X$, in which case $x \prec y$.

It remains just to show that $<$ is a subset of $\prec$. Suppose that $x < y$. If $x, y \in A'$, then $x \prec' y$, hence $x \prec y$. Suppose that $x = a_0$. If $y \in X$, choose v so that $v < a_0$ and $y \preceq' v$. Then $v < a_0 < y$, so $v < y$, hence $v \prec' y$, contradiction. So $y \notin X$, and hence $x \prec y$. Finally, suppose that $y = a_0$. Then $x \in X$ by definition of X, so $x \prec y$.

This finishes the case when A is finite. Note that all of this is just standard set theory, no logic involved.

Now we take a partial ordering structure $(A, <)$ with A infinite. Let $\mathscr{L}$ be a first-order language which has a binary relation symbol $<$ and, for each $a \in A$, an individual constant $\mathbf{c}_a$. In this language we consider the following set Γ of sentences:

$$\begin{aligned} &\text{Sentences saying that } < \text{ gives a simple order} \\ &\mathbf{c}_a < \mathbf{c}_b \quad \text{whenever } a < b, \text{ for all } a, b \in A. \end{aligned}$$

By our previous argument, every finite subset of Γ has a model. So Γ itself has a model $(B, <, b_a)_{a \in A}$, where b_a is the denotation of $\mathbf{c}_a$ in the model. Now we define $a \prec a'$ iff $a, a' \in A$ and $b_a < b_{a'}$. Clearly $(A, \prec)$ is a simple ordering structure. If $a, a' \in A$ and $a < a'$, then the sentence $\mathbf{c}_a < \mathbf{c}_{a'}$ is in Γ, and so this sentence holds in $(B, <, b_a)_{a \in A}$, which means that $b_a < b_{a'}$ and hence $a \prec a'$. □

The following concept will play an important role in what follows. Two structures $\overline{M}, \overline{N}$ are *elementarily equivalent* iff for every sentence φ, $\overline{M} \models \varphi$ iff $\overline{N} \models \varphi$.

EXERCISES

Exc. 5.1. Suppose that $\overline{A}$ is an $\mathscr{L}$-structure. Let F be a nonprincipal ultrafilter on a set I. For each $a \in A$ let $f(a) = [\langle a : i \in I \rangle]_F$. Show that f is an embedding of $\overline{A}$ into ${}^I\overline{A}/F$, and $\overline{A}$ is elementarily equivalent to ${}^I\overline{A}/F$.

Exc. 5.2. We work in the language for ordered fields; see Chapter 1. In general, an element $a \in M$ is *definable* iff there is a formula $\varphi(x)$ with one free variable x such that $\{b \in M : \overline{M} \models \varphi[b]\} = \{a\}$.

(i) Show that 1 is definable in $\mathbb{R}$.

(ii) Show that every positive integer is definable in $\mathbb{R}$.

(iii) Show that every positive rational is definable in $\mathbb{R}$.

(iv) If $\overline{M}$ is an extension of $\mathbb{R}$, an element ε of M is *infinitesimal* iff $0 < \varepsilon < r$ for every positive rational r. Let $\overline{M}$ be an ultrapower of $\mathbb{R}$ using a nonprincipal ultrafilter on ω. Thus $\overline{M}$ is isomorphic to an extension of $\mathbb{R}$ by exercise 5.1. Show that $\overline{M}$ has an infinitesimal.

(v) Use the compactness theorem to show the existence of an ordered field $\overline{M}$ which has an infinitesimal and is elementarily equivalent to $\mathbb{R}$.

Exc. 5.3. Consider the structure $\overline{N} = (\omega, +, \cdot, 0, 1, <)$. We look at models of $\Gamma = \{\varphi : \varphi$ is a sentence and $\overline{N} \models \varphi\}$.

(i) For every $m \in \omega$ there is a formula φ_m with one free variable x such that $\overline{N} \models \varphi_m[m]$ and $\overline{N} \models \exists! x \varphi_m(x)$.

(ii) $\overline{N}$ can be embedded in any model of Γ.

(iii) Show that Γ has a model with an infinite element in it, i.e., an element greater than each $m \in \omega$.

Exc. 5.4. (Continuing exercise 5.3.) An element p of a model $\overline{M}$ of Γ is a *prime* iff $p > 1$ and for all $a, b \in M$, if $p = a \cdot b$ then $a = 1$ or $a = p$.

(i) Prove that if $\overline{M}$ is a model of Γ with an infinite element, then it has an infinite prime element.

(ii) Show that the following conditions are equivalent:

(a) There are infinitely many (ordinary) primes p such that $p + 2$ is also prime. (The famous twin prime conjecture, unresolved at present.)

(b) There is a model $\overline{M}$ of Γ having at least one infinite prime p such that $p + 2$ is also a prime.

(c) For every model $\overline{M}$ of Γ having an infinite element, there is an infinite prime p such that $p + 2$ is also a prime.

Exc. 5.5. Let G be a group which has elements of arbitrarily large finite order. Show that there is a group H elementarily equivalent to G which has an element of infinite order.

Exc. 5.6. Suppose that Γ is a set of sentences, and φ is a sentence. Prove that if $\Gamma \models \varphi$, then $\Delta \models \varphi$ for some finite $\Delta \subseteq \Gamma$.